# On LAN for parametrized continuous periodic signals in a time inhomogeneous diffusion

**R. Höpfner**, Johannes Gutenberg Universität Mainz

**Yu. Kutoyants**, Université du Maine, Le Mans

**Abstract:** We consider a diffusion $(\xi_t)_{t\geq 0}$ whose drift involves some $T$-periodic signal. $T$ is fixed and known, whereas the signal depends on a $d$-dimensional parameter $\vartheta \in \Theta$. Assuming positive Harris recurrence of the grid chain $(\xi_{kT})_{k\in I\!\!N_0}$ and exploiting the periodic structure in the semigroup, we work with path segments and limit theorems for certain functionals (more general than additive functionals) of the process to prove local asymptotic normality (LAN). Then we consider several estimators for the unknown parameter.

**Key words:** diffusions, inhomogeneity in time, periodicity, continuous signals, local asymptotic normality, local asymptotic minimx bound, minimum distance estimators, one-step correction.

**MSC:** **62 F 12 , 60 J 60**

March 19, 2010

We consider a problem of parameter estimation in a time-inhomogenous diffusion $(\xi_t)_{t\geq 0}$ whose semi-group has a $T$-periodic structure, and whose drift involves a signal –deterministic, continuous, of known periodicity $T$– which is parametrized by some $d$-dimensional parameter $\vartheta \in \Theta$. Our main assumption on the process is positive Harris recurrence of the grid chain $(\xi_{kT})_{k\in I\!\!N_0}$ which implies (as in [HK 10]) positive Harris recurrence of the path-segment chain

$$\mathbb{X} = (\mathbb{X}_k)_k \quad , \quad \mathbb{X}_k := \big(\xi_{(k-1)T+s}\big)_{0\leq s\leq T}$$

taking values in $C([0,T])$: this allows for limit theorems using the path-segment chain. Our main statistical assumption is that the derivative of the signal with respect to the parameter is again a $T$-periodic function, and that the parametrization is sufficiently smooth, in a suitable $L^2$-sense related to the periodic structure of the semigroup.

We prove local asymptotic normality (LAN) for the sequence of statistical models corresponding to observation of a trajectory of $\xi$ – with unknown $\vartheta$ – over a long time interval $[0, nT]$, and state a local asymptotic minimax theorem. See LeCam ([LC 68], [LC 90]), Hájek [H 71], Ibragimov and

Khasminskii [IH 81] for background on LAN and convergence of estimators, see also the survey paper on local asymptotics by Davies [D 85], and see Liptser and Shiryaev [LS 81], Jacod and Shiryaev [JS 87], Kutoyants [K 04] on background on statistical models and likelihood ratio processes in diffusions. 'Signal in white noise' problems seem to appear first in Ibragimov and Khasminskii [IH 81]. The limit theorems which we use for time-inhomogeneous diffusions $(\xi_t)_{t \geq 0}$ with periodic structure in the semigroup are from [HK 10, Section 2].

We then consider a sequence of minimum distance estimators (MDE) which via LeCam's one step correction can be improved to get an asymptotically efficient estimator sequence, thus providing an easy-to-calculate and explicit alternative to maximum likelihood estimators and Bayes estimators which under some conditions are asymptotically efficient for the unknown parameter.

## 1. Local asymptotic normality

Our setting is as follows. The observed diffusion is inhomogeneous in time

$$(1) \qquad d\xi_t \;=\; [S(\vartheta, t) + b(\xi_t)]\, dt \;+\; \sigma(\xi_t)\, dW_t \quad,\quad t \geq 0\,,\; \xi_0 = x_0$$

and depends on an unknown parameter $\vartheta$ ranging over an open set $\Theta \subset I\!\!R^d$. The drift in (1) involves a deterministic signal $S(\cdot, \cdot)$, which is a continuous function $\Theta \times [0, \infty) \to I\!\!R$ satisfying

$$(2) \qquad \text{for every } \vartheta \in \Theta \text{ fixed:} \quad S(\vartheta, \cdot) : [0, \infty) \to I\!\!R \quad \text{is } T\text{-periodic}$$

for some period $T$ (fixed and known, and not depending on $\vartheta$). We write throughout the paper

$$i_T(t) \;:=\; t \bmod T\,.$$

The functions $b(\cdot)$ and $\sigma(\cdot)$ are assumed Lipschitz, and $\sigma(\cdot)$ strictly positive. Hence for all values of $\vartheta \in \Theta$, we have Lipschitz and linear growth conditions for the time-dependent coefficients of equation (1), and thus existence and pathwise uniqueness for its solution ([S 65], [KS 91]). We discuss 'ergodicity properties' for the time inhomogeneous diffusion $\xi$. Write $(P_{s,t}^{(\vartheta)})_{0 \leq s < t < \infty}$ for the semigroup of the process (1). As a consequence of (2), this semigroup is $T$-periodic in the sense

$$(3) \qquad P_{s,t}^{(\vartheta)}(x, dy) \;=\; P_{s+kT, t+kT}^{(\vartheta)}(x, dy) \quad \text{for all } k \in I\!\!N_0 \text{ and all } 0 \leq s < t < \infty\,.$$

We assume

$$(H1) \qquad \begin{array}{l} \text{for every } \vartheta \in \Theta, \text{ the grid chain } (\xi_{kT})_{k \in I\!\!N_0} \text{ is positive recurrent} \\[4pt] \text{in the sense of Harris with invariant probability } \mu^{(\vartheta)} \text{ on } (I\!\!R, \mathcal{B}(I\!\!R)) \end{array}$$

and decompose the path of $\xi$ into a Markov chain of $T$–segments

$$\mathbb{X} = (\mathbb{X}_k)_k \quad \text{defined by} \quad \mathbb{X}_k := (\xi_{(k-1)T+s})_{0 \le s \le T} \, , \; k \ge 1$$

which takes values in the path space $(C_T, \mathcal{C}_T)$ of continuous functions $[0,T] \to I\!\!R$, using as initial value $\mathbb{X}_0$ some $\alpha \in C_T$ such that $\alpha(T) = x_0$. As a consequence of (3), the chain $\mathbb{X}$ is time homogeneous. As a consequence of $(H1)$, $\mathbb{X}$ will be positive recurrent in the sense of Harris under $\vartheta$ (see [HK 09, theorem 2.1]), with invariant probability $m^{(\vartheta)}$ on $(C_T, \mathcal{C}_T)$ whose finite-dimensional distributions are given in terms of $\mu^{(\vartheta)}$ and $P_{s,t}^{(\vartheta)}$, $0 \le s < t \le T$. In order to obtain limit theorems for log-likelihoods related to observation of the time-inhomogeneous process $\xi$ over a long time interval, for asymptotics of local models at $\vartheta$, or for rescaled estimation errors of interesting estimators, we shall always exploit the ergodicity of the time homogeneous $C_T$-valued chain $\mathbb{X}$.

We turn to properties of the parametrization and list the properties of the parametrization which we shall need. Our identifiability condition is

$(H2)$ $\qquad\qquad$ for all $\vartheta \in \Theta$, $\varepsilon > 0$: $\quad \inf\limits_{|\zeta - \vartheta| > \varepsilon} \sup\limits_{0 \le s \le T} |S(\zeta, s) - S(\vartheta, s)| > 0$ .

A sufficiently general differentiability condition is as follows. Restrict the continuous periodic functions $\{S(\zeta, \cdot) : \zeta \in \Theta\}$ to $[0,T]$, assume that all measures

$$(4) \qquad\qquad \lambda^{(\vartheta)}(ds) := \left[ \mu^{(\vartheta)} P_{0,s}^{(\vartheta)}(\frac{1}{\sigma^2}) \right] ds \quad \text{on} \; ([0,T], \mathcal{B}([0,T])) \quad , \quad \vartheta \in \Theta$$

are finite, and consider spaces

$$\mathbb{H}^{(\vartheta)} = L^2([0,T], \mathcal{B}([0,T]), \lambda^{(\vartheta)}) \quad , \quad \vartheta \in \Theta \; .$$

Assume that for every $\vartheta \in \Theta$, there is a function

$$\dot{S}(\vartheta, \cdot) = \begin{pmatrix} \dot{S}_1(\vartheta, \cdot) \\ \dots \\ \dot{S}_d(\vartheta, \cdot) \end{pmatrix} \quad , \quad \dot{S}_j(\vartheta, \cdot) \in \mathbb{H}^{(\vartheta)} \; \text{for} \; j = 1, \dots, d$$

such that the following holds:

$(H3)$ $\quad \begin{cases} \rho_\varepsilon(\vartheta, \cdot) := \sup\limits_{|\zeta - \vartheta| < \varepsilon} \left| \frac{S(\zeta, \cdot) - S(\vartheta, \cdot) - (\zeta - \vartheta)^\top \dot{S}(\vartheta, \cdot)}{\zeta - \vartheta} \right| \in \mathbb{H}^{(\vartheta)} \quad \text{for some} \; \varepsilon = \varepsilon(\vartheta) > 0 \\ \rho_\varepsilon(\vartheta, \cdot) \longrightarrow 0 \quad \text{in} \; \mathbb{H}^{(\vartheta)} \; \text{as} \; \varepsilon \downarrow 0 \; . \end{cases}$

This is slightly more than Fréchét differentiability of $\zeta \to S(\zeta, \cdot)$ at $\zeta = \vartheta$ in the Hilbert space $\mathbb{H}^{(\vartheta)}$. The assumption that the derivative with respect to the parameter $t \to \dot{S}(\vartheta, t)$ is again a $T$-periodic

function –the key to the results presented here– is a strong assumption; e.g., it rules out estimation of an unknown periodicity ('frequency modulation') as considered in [IH 81, p. 209] or [CLM 06]. We call the $d \times d$ matrix

$$(5) \qquad I^{(\vartheta)} := \int_0^T \dot{S}(\vartheta, s)\, \dot{S}(\vartheta, s)^\top \lambda^{(\vartheta)}(ds)$$

Fisher information at $\vartheta$, and add the assumption

$$(H4) \qquad\qquad I^{(\vartheta)} \text{ is invertible for all } \vartheta \in \Theta \,.$$

We introduce two additional assumptions. For $S := \{\alpha 2^{-k} : k \in \mathbb{N}_0\,, \alpha \in \mathbb{Z}^d\}$ the set of dyadic numbers in $\mathbb{R}^d$, we shall need

$$(H5) \qquad \begin{cases} \widetilde{\rho}_\varepsilon(\vartheta, \cdot) := \sup\limits_{\zeta \in S \cap \Theta, |\zeta - \vartheta| < \varepsilon} \left| \dot{S}(\zeta, \cdot) - \dot{S}(\vartheta, \cdot) \right| \;\; \in \;\; \mathbb{H}^{(\vartheta)} \quad \text{for some } \varepsilon = \varepsilon(\vartheta) > 0 \\[2mm] \widetilde{\rho}_\varepsilon(\vartheta, \cdot) \;\longrightarrow\; 0 \quad \text{in } \mathbb{H}^{(\vartheta)} \;\; \text{as } \varepsilon \downarrow 0 \end{cases}$$

for all $\vartheta \in \Theta$ together with

$$(H6) \qquad\qquad \sup_{\zeta \in S \cap \Theta\,, |\zeta - \vartheta| < \varepsilon} \left| I^{(\zeta)} - I^{(\vartheta)} \right| \;\longrightarrow\; 0 \quad \text{as } \;\; \varepsilon \downarrow 0 \,.$$

Let $Q^\vartheta$ denote the law of the solution of (1)+(2) under $\vartheta \in \Theta$, a law on the canonical path space $(C, \mathcal{C})$ of continuous functions $[0, \infty) \to \mathbb{R}$ equipped with the metric of locally uniform convergence. Write $Q_t^\vartheta$ for the restriction of $Q^\vartheta$ to the $\sigma$-field $\mathcal{G}_t$ of events before time $t^+$, and $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$. The likelihood process $L^{\zeta/\vartheta}$ of $Q^\zeta$ with respect to $Q^\vartheta$ relative to $\mathbb{G}$ (cf. [LS 81], [JS 87], [K 04]) is given by

$$(6) \qquad L^{\zeta/\vartheta} \;=\; \mathcal{E}_\vartheta\left( \int_0^\cdot \frac{S(\zeta, t) - S(\theta, t)}{\sigma^2(\eta_t)} \, dm_t^{(\vartheta)} \right) \;=\; \mathcal{E}_\vartheta\left( \int_0^\cdot \frac{S(\zeta, t) - S(\theta, t)}{\sigma(\eta_t)} \, dB_t^{(\vartheta)} \right)$$

where $\eta = (\eta_t)_{t \geq 0}$ denotes the canonical process on $(C, \mathcal{C}, \mathbb{G})$, $m^{(\vartheta)}$ its $(\mathbb{G}, Q^\vartheta)$-martingale part, and where $B^{(\vartheta)} := \int_0^\cdot \frac{1}{\sigma(\eta_s)} dm_s^{(\vartheta)}$ is a $(Q^\vartheta, \mathbb{G})$-Brownian motion. The following is a '2nd Le Cam lemma' for time inhomogeneous diffusions of type (1)+(2):

**1.1 Theorem :** Under $(H1)+(H3)$, in the sequence of experiments

$$\left( C\,, \mathcal{G}_{nT}\,, \left\{ Q_{nT}^\zeta : \zeta \in \Theta \right\} \right) \quad,\quad n \to \infty\,,$$

we have LAN at every point $\vartheta \in \Theta$, with local scale $n^{-1/2}$ and Fisher information $I^{(\vartheta)}$ given by (5). More precisely, for every $\vartheta \in \Theta$ and arbitrary bounded sequences $(h_n)_n$ in $\mathbb{R}^d$, we have a quadratic decomposition of log-likelihood ratios

$$(7) \qquad\qquad \log L_{nT}^{(\vartheta + n^{-1/2}h_n)/\vartheta} \;=\; h_n^\top \Delta_n^{(\vartheta)} \;-\; \frac{1}{2} h_n^\top I^{(\vartheta)} h_n \;+\; o_{Q^\vartheta}(1) \quad \text{as } n \to \infty$$

with score

$$\Delta_n^{(\vartheta)} = \frac{1}{\sqrt{n}} \int_0^{nT} \frac{\dot{S}(\vartheta, s)}{\sigma(\eta_s)} \, dB_s^{(\vartheta)}$$

such that

(8) $$\qquad \mathcal{L}(\Delta_n^{(\vartheta)} \mid Q^\vartheta) \longrightarrow \mathcal{N}(0, I^{(\vartheta)}) \quad \text{(weak convergence in } {I\!\!R}^d, \, n \to \infty).$$

The proof of theorem 1.1 will be given in section 2 below (see 2.1). For background on LAN see [L 69], [H 70], [IH 81], [D 85], [LY 90], [K 04], or also [H 08] for some more details. A well-known and powerful consequence of LAN is the local asymptotic minimax theorem (corollary 1.1' below, from [L 90, p. 83]; a result using neighbourhoods in a different way is [IH 81], with additional uniformity assumptions which we do not make here). Fix any reference point $\vartheta \in \Theta$, write

(9) $$\qquad Z_n^{(\vartheta)} := (I^{(\vartheta)})^{-1} \Delta_n^{(\vartheta)} \;, \quad n \ge 1$$

for the central sequence in the sequence of local models

$$\mathcal{E}_n(\vartheta) := \left\{ P_{n,\vartheta,h} := Q_{nT}^{\vartheta + n^{-1/2}h} \; : \; h \in {I\!\!R}^d \text{ such that } \vartheta + n^{-1/2}h \in \Theta \right\} \;, \quad n \ge 1$$

at $\vartheta$, and let $\ell(\cdot)$ denote any loss function which is continuous, subconvex and bounded. Then

**1.1' Corollary :** (LeCam [L 90, p. 83]) Under $(H1)$ and $(H3)+(H4)$, the following holds. For any sequence of $\mathcal{G}_{nT}$-measurable estimators $\widetilde{\vartheta}_n$, for any $\vartheta \in \Theta$,

$$\sup_{c<\infty} \liminf_{n\to\infty} \sup_{|h|\le c} E_{(\vartheta+n^{-1/2}h)} \left( \ell \left( n^{1/2}(\widetilde{\vartheta}_n - (\vartheta + n^{-1/2}h)) \right) \right) \;\ge\; \int_{-\infty}^{\infty} \ell(z) \, \mathcal{N}(0, (I^{(\vartheta)})^{-1})(dz).$$

Sequences of $\mathcal{G}_{nT}$-measurable estimators $\vartheta_n^*$ with the property

(10) $$\qquad n^{1/2}(\vartheta_n^* - \vartheta) = Z_n^{(\vartheta)} + o_{Q^\vartheta}(1) \quad \text{as } n \to \infty$$

attain the local asymptotic minimax bound at $\vartheta$, and have

$$\liminf_{n\to\infty} \sup_{|h|\le c} E_{(\vartheta+n^{-1/2}h)} \left( \ell \left( n^{1/2}(\vartheta_n^* - (\vartheta + n^{-1/2}h)) \right) \right) \;=\; \int_{-\infty}^{\infty} \ell(z) \, \mathcal{N}(0, (I^{(\vartheta)})^{-1})(dz)$$

for every $0 < c < \infty$.

Estimator sequences attaining the local asymptotic minimax bound at $\vartheta$ are termed efficient at $\vartheta$. The criterion (10) characterizes efficiency at $\vartheta$ in terms of stochastic equivalence under $\vartheta$ of rescaled

estimation errors at $\vartheta$ and central sequence at $\vartheta$. Using Le Cam's 'one step correction', efficient estimator sequences can be constructed explicitely provided we dispose of at least one preliminary estimator sequence which is $n^{1/2}$-consistent. The construction which we give here follows Davies [D 85].

**Theorem 1.2 :** Assume $(H1)$, $(H3)+(H4)$, $(H5)+(H6)$, and let $\widetilde{\vartheta}_n$ denote some sequence of $\mathcal{G}_{nT}$-measurable estimators with the property

$$\text{for every } \vartheta \in \Theta : \quad \mathcal{L}\left(n^{1/2}(\widetilde{\vartheta}_n - \vartheta) \mid Q^{(\vartheta)}\right) , \ n \geq 1 , \text{ is tight in } I\!\!R^d .$$

For $\alpha = (\alpha_1, \ldots, \alpha_d) \in Z\!\!\!Z^d$ and $k \in I\!\!N_0$, consider cubes $C(k,\alpha) := \underset{j=1}{\overset{d}{\mathsf{X}}} [\alpha_j 2^{-k}, (\alpha_j+1)2^{-k}[$ in $I\!\!R^d$. Fixing some point $\vartheta_0 \in \Theta$, we discretize the estimator $\widetilde{\vartheta}_n$ to

$$\widetilde{\widetilde{\vartheta}}_n := \vartheta_0 + \sum_{\alpha \in Z\!\!\!Z^d \text{ s.t. } C(n,\alpha) \subset \Theta} (\alpha 2^{-n} - \vartheta_0) \, 1_{C(n,\alpha)}(\widetilde{\vartheta}_n)$$

which takes a countable number of values in $\Theta$. Then

$$\vartheta_n^* := \widetilde{\widetilde{\vartheta}}_n + n^{-1/2} \, (I^{(\widetilde{\widetilde{\vartheta}}_n)})^{-1} \, \Delta_n^{(\widetilde{\widetilde{\vartheta}}_n)}$$

is a sequence of $\mathcal{G}_{nT}$-measurable estimators for $\vartheta \in \Theta$ which has the property (10) at all points $\vartheta \in \Theta$.

Thus by corollary 1.1', the sequence $(\vartheta_n^*)_n$ constructed in theorem 1.2 is efficient at all points $\vartheta \in \Theta$. Since sequences meeting (10) are regular at $\vartheta$ in the sense of Hájek ([H 70]), they are also regular and efficient in the sense of Hájek's convolution theorem.

**1.3 Example:** We look to equations (1) of Ornstein-Uhlenbeck type

$$b(x) = -\beta x \quad \text{for some } \beta > 0 \quad , \quad \sigma(\cdot) \equiv \sigma > 0 \quad \text{constant} .$$

1) It follows from [HK 10, example 2.3] that $(H1)$ holds for arbitrary continuous and $T$-periodic functions $t \to S(\vartheta, t)$, $\vartheta \in \Theta$. The diffusion coefficient being constant, the measure $\lambda^{(\vartheta)}$ in (4) is Lebesgue measure on $[0, T]$ multiplied by the constant $\frac{1}{\sigma^2}$. It does not depend on $\vartheta \in \Theta$.

2) Let $\Theta \subset I\!\!R^d$ be open, let $(K_m)_m$ denote an increasing sequence of compacts with $K_m \subset int(K_{m+1})$, $m \geq 1$, and $\Theta = \bigcup_m K_m$. For every $t \in [0, T]$ fixed, let the mapping $\zeta \to S(\zeta, t)$ be $\mathcal{C}^1$ on $\Theta$. Then

(11) $$\text{for every } m : \quad \sup_{t \in [0,T]} \sup_{\zeta \in K_m} \sup_{|u|=1} \left| u^\top \dot{S}(\zeta, t) \right| \quad < \quad \infty$$

is a sufficient condition implying $(H3)+(H5)+(H6)$. Under (11), at every point $\vartheta \in \Theta$,

$$\sup_{\zeta \in K_m} \frac{1}{|\zeta - \vartheta|^2} \int_0^T (S(\zeta, t) - S(\vartheta, t))^2 \, dt \quad < \quad \infty$$

provided $m$ is large enough for $\vartheta \in int(K_m)$, and we may consider the following partially converse condition: for all $\vartheta$, assume that there is some $\varepsilon = \varepsilon(\vartheta) > 0$ such that

$$(12) \qquad \inf_{\zeta \in \Theta} \frac{1}{|\zeta - \vartheta|^2 \wedge \varepsilon} \int_0^T (S(\zeta, t) - S(\vartheta, t))^2 \, dt \quad > \quad 0 \, .$$

Then under (11), condition (12) implies both $(H2)+(H4)$ together. Hence, if $\zeta \to S(\zeta, t)$ is $\mathcal{C}^1$ on $\Theta$ for fixed $t$, $(11)+(12)$ are sufficient conditions for $(H2)$-$(H6)$.

3) Fix some continuous $T$-periodic function $f(\cdot)$, non-constant. Put $\Theta := (0, T) \times (0, \infty)$, write $\vartheta = (\vartheta_1, \vartheta_2)$, and define (with notation $i_T(t)$ for $t$ modulo $T$ as above)

$$S(\vartheta, t) := \vartheta_2 \, f(i_T(t) - \vartheta_1) \quad , \quad t \in \mathbb{R} \, .$$

Then $(11)+(12)$ and thus $(H2)$-$(H6)$ hold. By $T$-periodicity of $f(\cdot)$, the Fisher information in (5) is

$$I^{(\vartheta)} \;=\; \frac{1}{\sigma^2} \begin{pmatrix} \vartheta_2^2 \int_0^T [f']^2(s)ds & -\vartheta_2 \int_0^T [f'f](s)ds \\ -\vartheta_2 \int_0^T [f'f](s)ds & \int_0^T f^2(s)ds \end{pmatrix}$$

which does not depend on $\vartheta_1 \in (0, T)$.

4) Assume that $T$ is large compared to 2, put $\Theta = (0, T)$, and define

$$S(\vartheta, t) \;:=\; f(t - \vartheta) \quad \text{with} \quad f(x) = 1_{\{|i_T(t)| \leq 1\}} (1 - |i_T(t)|) \, , \; t \in \mathbb{R} \, .$$

Then assumptions $(H2) - (H6)$ are satisfied with

$$\dot{S}(\vartheta, t) \;=\; \left( 1_{(\vartheta, \vartheta+1)} - 1_{(\vartheta-1, \vartheta)} \right) (i_T(t)) \quad , \quad t \in \mathbb{R}$$

and Fisher information

$$I^{(\vartheta)} \;=\; \frac{2}{\sigma^2} \quad , \quad \vartheta \in \Theta$$

not depending on $\vartheta \in \Theta$. Similiarly, if we replace the definition of $f$ above by

$$f(x) = 1_{\{|i_T(t)| \leq 1\}} (1 - |i_T(t)|)^\alpha \, , \; t \in \mathbb{R}$$

with some $\alpha > 1$, then assumptions $(H2) - (H6)$ are satisfied: we have

$$\dot{S}(\vartheta, t) \;=\; g(t - \vartheta) \quad \text{with} \quad g(t) := \alpha \, (1 - |i_T(t)|)^{\alpha-1} \left( 1_{(0,1)} - 1_{(-1,0)} \right) (i_T(t))$$

and the Fisher information does not depend on $\vartheta \in \Theta$. This is analogous to some examples in [L 90, p. 32]: $\dot{S}(\vartheta, \cdot)$ as element of $\mathbb{H}^{(\vartheta)}$ is defined for $\alpha > \frac{1}{2}$, $(H3)$ holds for $\alpha > \frac{1}{2}$ by dominated convergence, but $(H5)$ is violated in case $1 > \alpha > \frac{1}{2}$. $\qquad\qquad \square$

## 2. Proofs

In this section, we assume that the strong Markov process $\xi$ lives on some $(\Omega, \mathcal{A}, \mathbb{F}, (P_x)_{x \in \mathbb{R}})$; 'almost surely' means almost surely with respect to every $P_x$, $x \in \mathbb{R}$; $\mathbb{F}^\xi$ is the filtration generated by $\xi$, and an $\mathbb{F}^\xi$–increasing process is an $\mathbb{F}^\xi$–adapted càdlàg process $A = (A_t)_{t \geq 0}$ with nondecreasing paths and $A_0 = 0$, almost surely.

$(C_T, \mathcal{C}_T)$ – the space of all continuous functions $\alpha : [0, T] \to \mathbb{R}$ with Borel $\sigma$-field $\mathcal{C}_T$ – is a Polish space, and $\mathcal{C}_T$ is also generated by the coordinate projections $\pi_t$, $0 \leq t \leq T$. The one-step-transition kernel $Q(\cdot, \cdot)$ of the chain $\mathbb{X}$ of $T$-segments in the path of $\xi$ is given by

$$Q(\alpha, F) \; := \; P\left( (\xi_s)_{0 \leq s \leq T} \in F \mid \xi_0 = \alpha(T) \right) \;, \quad \alpha \in C_T \;, \; F \in \mathcal{C}_T \;.$$

For Harris processes in discrete time, we refer to Revuz [R 75] or Meyn and Tweedie [MT 93]. For Harris processes in continuous time see Azéma, Duflo and Revuz [ADR 69] or Révuz and Yor [RY 91, Ch. 10.3]. The following (with sub- or superscripts $\vartheta$ suppressed from notation) is a precise statement of the 'ergodicity properties' mentioned in section 1.

**Theorem A:** ( [HK 10, theorem 2.1]) Assume that the grid chain $(\xi_{kT})_{k \in \mathbb{N}_0}$ is positive recurrent in the sense of Harris, and write $\mu$ for its invariant probability on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

a) Then the chain $\mathbb{X} = (\mathbb{X}_k)_{k \in \mathbb{N}_0}$ of $T$-segments in the path of $\xi$ is positive recurrent in the sense of Harris. Its invariant probability is the unique law $m$ on $(C_T, \mathcal{C}_T)$ such that

(13)
$$\begin{cases} \text{for arbitrary } 0 = t_0 < t_1 < \ldots < t_\ell < t_{\ell+1} = T \text{ and } A_i \in \mathcal{B}(\mathbb{R}) \;, \\ m\left( \{ \pi_{t_i} \in A_i \,, \, 0 \leq i \leq \ell+1 \} \right) \quad \text{is given by} \\ \int \ldots \int \mu(dx_0) \, 1_{A_0}(x_0) \prod_{i=0}^{\ell} P_{t_i, t_{i+1}}(x_i, dx_{i+1}) \, 1_{A_{i+1}}(x_{i+1}) \;. \end{cases}$$

b) For every $\mathbb{F}^\xi$–increasing process $A = (A_t)_{t \geq 0}$ with the property

(14)
$$\begin{cases} \text{there is some function } F : C_T \to \mathbb{R}, \text{ nonnegative, } \mathcal{C}_T\text{-measurable,} \\ \text{satisfying} \quad m(F) := \int_{C_T} F(\alpha) \, m(d\alpha) \; < \; \infty \;, \quad \text{such that} \\ A_{kT} \; = \; \sum_{j=1}^{k} F(\mathbb{X}_k) \; = \; \sum_{j=1}^{k} F\left( (\xi_{(k-1)T+s})_{0 \leq s \leq T} \right) \;, \; k \geq 1 \end{cases}$$

we have the strong law of large numbers

$$\lim_{t \to \infty} \frac{1}{t} A_t \; = \; \frac{1}{T} m(F) \quad \text{almost surely} \;.$$

8

This strong law of large numbers will be the key tool to prove theorem 1.1. We use the notations which have been introduced in section 1.

**2.1 Proof of theorem 1.1 :**   We assume $(H1)$ and $(H3)$.

1)  Fix $\vartheta$. For $h \in I\!\!R^d$, consider

$$\log L_{nT}^{(\vartheta + n^{-1/2}h)/\vartheta} \;=\; \int_0^{nT} \frac{S(\vartheta + n^{-1/2}h, s) - S(\vartheta, s)}{\sigma(\eta_s)} \, dB_s^{(\vartheta)} - \frac{1}{2} \int_0^{nT} \left( \frac{S(\vartheta + n^{-1/2}h, s) - S(\vartheta, s)}{\sigma(\eta_s)} \right)^2 ds$$

under $Q^\vartheta$. Adding $\pm n^{-1/2} h^\top \dot{S}(\vartheta, \cdot)$ in the numerators of the integrands, we separate leading terms

$$(15) \hspace{3cm} h^\top \Delta_T^n(\vartheta) \;-\; \frac{1}{2} h^\top I_T^n(\vartheta) \, h$$

defined by

$$(16) \hspace{3cm} \Delta_t^n(\vartheta) \;:=\; \frac{1}{\sqrt{n}} \int_0^{tn} \frac{\dot{S}(\vartheta, s)}{\sigma(\eta_s)} \, dB_s^{(\vartheta)} \;\;,\;\; t \geq 0$$

$$(17) \hspace{3cm} I_t^n(\vartheta) \;:=\; \frac{1}{n} \int_0^{tn} \left( \frac{\dot{S}(\vartheta, s)\, \dot{S}(\vartheta, s)^\top}{\sigma^2(\eta_s)} \right) ds \;\;,\;\; t \geq 0$$

from remainder terms

$$(18) \hspace{3cm} N_T^n(h) \;-\; \frac{1}{2} U_T^n(h) \;-\; V_T^n(h)$$

defined by

$$(19) \hspace{1.5cm} N_t^n(h) \;:=\; \int_0^{tn} \frac{S(\vartheta + n^{-1/2}h, s) - S(\vartheta, s) - n^{-1/2} h^\top \dot{S}(\vartheta, s)}{\sigma(\eta_s)} \, dB_s^{(\vartheta)} \;\;,\;\; t \geq 0$$

$$(20) \hspace{1.5cm} U_t^n(h) \;:=\; \int_0^{tn} \left( \frac{S(\vartheta + n^{-1/2}h, s) - S(\vartheta, s) - n^{-1/2} h^\top \dot{S}(\vartheta, s)}{\sigma(\eta_s)} \right)^2 ds$$

$$(21) \hspace{1cm} V_t^n(h) \;:=\; \int_0^{tn} \frac{[S(\vartheta + n^{-1/2}h, s) - S(\vartheta, s) - n^{-1/2} h^\top \dot{S}(\vartheta, s)][n^{-1/2} h^\top \dot{S}(\vartheta, s)]}{\sigma^2(\eta_s)} \, ds \;.$$

We shall show convergence of $I_T^n(\vartheta)$ in (17) to the Fisher information

$$I^{(\vartheta)} \;=\; \int_0^T \dot{S}(\vartheta, s)\, \dot{S}(\vartheta, s)^\top \lambda^{(\vartheta)}(ds)$$

$Q^\vartheta$-almost surely as $n \to \infty$, weak convergence of $\mathcal{L}\left( \Delta_T^n(\vartheta) \mid Q^\vartheta \right)$ in (16) to $\mathcal{N}(0, I^{(\vartheta)})$, as well as

$$(22) \hspace{2cm} N_T^n(h_n)\,,\, U_T^n(h_n)\,,\, V_T^n(h_n) \;\;\longrightarrow\;\; 0 \;\; \text{in } Q^\vartheta\text{-probability as } n \to \infty$$

for arbitrary bounded sequences $(h_n)_n$ in $I\!R^d$.

2) Consider first the processes $(I_t^n(\vartheta))_{t\geq 0}$ defined in (17). Write $F_{j,j'} : C_T \to I\!R$ for the function

$$\alpha \longrightarrow F_{j,j'}(\alpha) := \int_0^T \left( \frac{\dot{S}_j(\vartheta, s)\dot{S}_{j'}(\vartheta, s)}{\sigma^2(\alpha(s))} \right) ds$$

with arbitrary $1 \leq j, j' \leq d$. By definition of $\lambda^{(\vartheta)}(ds)$ in (4) and since $\dot{S}_j(\vartheta, \cdot) \in \mathbb{H}^{(\vartheta)}$ by $(H3)$,

$$(23) \qquad m^{(\vartheta)}(F_{j,j'}) \;=\; \int_{C_T} F_{j,j'}(\alpha)\, m^{(\vartheta)}(d\alpha) \;=\; \int_0^T \dot{S}_j(\vartheta, s)\dot{S}_{j'}(\vartheta, s)\, \lambda^{(\vartheta)}(ds)$$

is well defined and finite, and equals $\left[I^{(\vartheta)}\right]_{j,j'}$. Combining (17) with theorem A , we obtain

$$(24) \qquad [I_T^n(\vartheta)]_{j,\ell} \;=\; \frac{1}{n} \sum_{\ell=1}^n F_{j,j'}(\mathbb{X}_\ell) \;\longrightarrow\; m^{(\vartheta)}(F_{j,j'}) \;=\; \left[I^{(\vartheta)}\right]_{j,j'}$$

almost surely as $n \to \infty$. This is $Q^\vartheta$-almost sure convergence of $I_T^n(\vartheta)$ to the Fisher information $I^{(\vartheta)}$.

3) We consider the processes in (16): $(\Delta_t^n(\theta))_{t\geq 0}$ are martingales with angle brackett $(I_t^n(\theta))_{t\geq 0}$ relative to $Q^\vartheta$ and $(\mathcal{G}_{nt})_{t\geq 0}$. For any $\beta \in I\!R^d$ fixed, the function $F(\alpha) = \sum_{j,j'} \beta_j F_{j,j'}(\alpha)\beta_{j'}$ is nonnegative and $m^{(\vartheta)}$-integrable by step 1) above, we have

$$\frac{1}{n} \sum_{k=1}^{\lfloor \frac{tn}{T} \rfloor} F(\mathbb{X}_k) \;\leq\; \beta^\top I_t^n(\vartheta)\,\beta \;\leq\; \frac{1}{n} \sum_{k=1}^{\lceil \frac{tn}{T} \rceil} F(\mathbb{X}_k)$$

and thus by theorem A, for every $u \geq 0$ fixed,

$$(25) \qquad \lim_{n\to\infty} \beta^\top I_{uT}^n(\vartheta)\,\beta \;=\; u \cdot \int_0^T \left( \beta^\top \dot{S}(\vartheta, s) \right)^2 \lambda^{(\vartheta)}(ds) \;=\; u \cdot \beta^\top I^{(\vartheta)}\,\beta$$

almost surely as $n \to \infty$. From (25), the martingale theorem [JS 87, VIII.3.22] gives weak convergence in the Skorohod path space $\mathbb{D}$ as $n \to \infty$

$$\mathcal{L}\left( \left( \beta^\top \Delta_{uT}^n(\vartheta) \right)_{u\geq 0} \mid Q^\vartheta \right) \;\longrightarrow\; \mathcal{L}\left( \left( \beta^\top \widetilde{B}_u(\vartheta) \right)_{u\geq 0} \right)$$

where $\widetilde{B}(\vartheta)$ is $d$-dimensional Brownian motion with covariance matrix $I^{(\vartheta)}$. But $\beta \in I\!R^d$ was arbitrary. So we have in particular weak convergence in $I\!R^d$ as $n \to \infty$

$$\mathcal{L}\left( \Delta_T^n(\vartheta) \mid Q^\vartheta \right) \;\longrightarrow\; \mathcal{N}\left( 0, I^{(\vartheta)} \right) .$$

All assertions concerning convergence of the leading terms (15) are now proved.

4) We turn to the remainder terms in (18) and fix any bounded sequence $(h_n)_n$ in $I\!R^d$. We start with processes (20) and write, for $N$ arbitrarily large and fixed,

$$U_{NT}^n(h_n) \;=\; \sum_{k=1}^{nN} G(\mathbb{X}_k, n, h_n)$$

10

where $G(\cdot, n, h) : C_T \to [0, \infty)$ is the function

$$\alpha \longrightarrow G(\alpha, n, h) := \int_0^T \left( \frac{S(\vartheta + n^{-1/2}h, s) - S(\vartheta, s) - n^{-1/2}h^\top \dot{S}(\vartheta, s)}{\sigma(\alpha(s))} \right)^2 ds .$$

Define with $\rho_\varepsilon(\vartheta, \cdot) \in \mathbb{H}^{(\vartheta)}$ introduced in $(H3)$

$$G_\varepsilon(\alpha) := \int_0^T \left( \frac{\rho_\varepsilon(\vartheta, s)}{\sigma(\alpha(s))} \right)^2 ds , \quad \alpha \in C_T , \ \varepsilon > 0 .$$

For $(h_n)_n$ bounded by $c$, $(H3)$ allows to find for every $\varepsilon > 0$ some $n_0$ such that

(26) $$G(\alpha, n, h_n) \leq c^2 \, n^{-1} \, G_\varepsilon(\alpha) \quad \text{for all } \alpha \in C_T \text{ and all } n \geq n_0 .$$

For any $\varepsilon > 0$ fixed, the strong law of large numbers in theorem A combined with (26) shows that

$$\limsup_{n \to \infty} U_{NT}^n(h_n) \leq \lim_{n \to \infty} c^2 \frac{1}{n} \sum_{k=1}^{nN} G_\varepsilon(\mathbb{X}_k) = c^2 \, N \, m^{(\vartheta)}(G_\varepsilon)$$

almost surely. But the second part of assumption $(H3)$ shows that

$$m^{(\vartheta)}(G_\varepsilon) = \int_{C_T} G_\varepsilon(\alpha) \, m^{(\vartheta)}(d\alpha) = \int_0^T \rho_\varepsilon^2(\vartheta, s) \, \lambda^{(\vartheta)}(ds) \longrightarrow 0$$

as $\varepsilon$ decreases to 0. Both last assertions combined give

(27) $$\lim_{n \to \infty} U_{NT}^n(h_n) = 0 \quad \text{almost surely, for arbitrary fixed } N \text{ fixed .}$$

The processes $(U_t^n(h_n))_{t \geq 0}$ being increasing, this implies the following property of the paths:

(28) $\quad U_\bullet^n(h_n)$ vanish uniformly on compact time intervals, $Q^\vartheta$-almost surely as $n \to \infty$ .

5) Next, $(N_t^n(h_n))_{t \geq 0}$ defined in (19) are martingales with angle brackett $(U_t^n(h_n))_{t \geq 0}$ relative to $Q^\vartheta$ and $(\mathcal{G}_{nt})_{t \geq 0}$. Hence for every $n$, the process

$$\left( \sup_{0 \leq s \leq t} |N_s^n(h_n)|^2 \right)_{t \geq 0}$$

is dominated in the sense of [JS 87, p. 35] by $(U_t^n(h_n))_{t \geq 0}$. As a consequence of step 2) this implies

(29) $\quad N_\bullet^n(h_n)$ vanish uniformly on compact time intervals, in $Q^\vartheta$-probability as $n \to \infty$ .

6) Finally, the reasoning of step 4) combined with Schwarz inequality and step 2) shows that

(30) $\quad V_\bullet^n(h_n)$ vanish uniformly on compact time intervals, $Q^\vartheta$-almost surely as $n \to \infty$ .

Hence all remainder terms in (22) vanish, and the proof of theorem 2.1 is complete. $\qquad \square$

**2.2 Proof of theorem 1.2 :** We assume $(H1)$, $(H3)+(H4)$, $(H5)+(H6)$. The proof follows the approach of Davies [D 85, p. 849].

1) In a first step, we shall show that for every $\vartheta \in \Theta$ and every $c < \infty$ we have approximations

$$(31) \qquad \sup_{\zeta \in S \cap \Theta,\, |\zeta - \vartheta| \le c\, n^{-1/2}} \left| \Delta_n^{(\zeta)} - \left( \Delta_n^{(\vartheta)} - I^{(\vartheta)}[n^{1/2}(\zeta - \vartheta)] \right) \right| \;=\; o_{Q^\vartheta}(1) \quad \text{as } n \to \infty$$

with $S := \{\alpha 2^{-k} : k \in \mathbb{N}_0,\, \alpha \in \mathbb{Z}^d\}$ the set of dyadic numbers.

From (6) and theorem 1.1 we have for points $\zeta \in S \cap \Theta$ such that $|\zeta - \vartheta| \le c\, n^{-1/2}$

$$
\begin{aligned}
\Delta_n^{(\zeta)} \;&=\; \frac{1}{\sqrt{n}} \int_0^{nT} \frac{\dot{S}(\zeta, s)}{\sigma^2(\eta_s)} \, dm_s^{(\zeta)} \\
(32) \qquad &=\; \frac{1}{\sqrt{n}} \int_0^{nT} \frac{\dot{S}(\zeta, s)}{\sigma^2(\eta_s)} \, dm_s^{(\vartheta)} \;-\; \frac{1}{\sqrt{n}} \int_0^{nT} \frac{\dot{S}(\zeta, s)}{\sigma^2(\eta_s)} \, (S(\zeta, s) - S(\vartheta, s)) \, ds
\end{aligned}
$$

according to Girsanov theorem. The $Q^\vartheta$-martingale term on the right hand side of (32) is

$$\frac{1}{\sqrt{n}} \int_0^{nT} \frac{\dot{S}(\zeta, s)}{\sigma^2(\eta_s)} \, dm_s^{(\vartheta)} \;=\; \Delta_n^{(\vartheta)} \;+\; o_{Q^\vartheta}(1)$$

as $n \to \infty$: this follows if we apply $(H5)$ and theorem A to angle bracketts of

$$\frac{1}{\sqrt{n}} \int_0^{nT} \frac{\dot{S}(\zeta, s) - \dot{S}(\vartheta, s)}{\sigma^2(\eta_s)} \, dm_s^{(\vartheta)} \quad, \quad n \to \infty \,.$$

Rewrite the bounded variation term on the right hand side of (32) as

$$
\begin{aligned}
&\frac{1}{n} \int_0^{nT} \frac{\dot{S}(\vartheta, s)\dot{S}(\vartheta, s)^\top}{\sigma^2(\eta_s)} \, [\sqrt{n}(\zeta - \vartheta)] \, ds \\
+\; &\frac{1}{n} \int_0^{nT} \frac{(\dot{S}(\zeta, s) - \dot{S}(\vartheta, s))\dot{S}(\vartheta, s)^\top}{\sigma^2(\eta_s)} \, [\sqrt{n}(\zeta - \vartheta)] \, ds \\
+\; &\frac{1}{n} \int_0^{nT} \frac{\dot{S}(\vartheta, s)}{\sigma^2(\eta_s)} \, |\sqrt{n}(\zeta - \vartheta)| \, \frac{S(\zeta, s) - S(\vartheta, s) - (\zeta - \vartheta)^\top \dot{S}(\vartheta, s)}{|\zeta - \vartheta|} \, ds \\
+\; &\frac{1}{n} \int_0^{nT} \frac{\dot{S}(\zeta, s) - \dot{S}(\vartheta, s)}{\sigma^2(\eta_s)} \, |\sqrt{n}(\zeta - \vartheta)| \, \frac{S(\zeta, s) - S(\vartheta, s) - (\zeta - \vartheta)^\top \dot{S}(\vartheta, s)}{|\zeta - \vartheta|} \, ds \,.
\end{aligned}
$$

The first summand equals

$$I^{(\vartheta)}\left[ \sqrt{n}(\zeta - \vartheta) \right] \;+\; o_{Q^\vartheta}(1)$$

by definition of the Fisher information and theorem A. The second summand, by Cauchy-Schwarz combined with $(H5)$, (5), theorem A and $|\zeta - \vartheta| \le c\, n^{-1/2}$, vanishes in $Q^\vartheta$-probability as $n \to \infty$. The third summand vanishes in $Q^\vartheta$-probability as $n \to \infty$ by Cauchy-Schwarz combined with $(H3)$, (5), theorem A and $|\zeta - \vartheta| \le c\, n^{-1/2}$. The forth summand behaves similiarly: we make use of $(H5)$, $(H3)$, theorem A and $|\zeta - \vartheta| \le c\, n^{-1/2}$. Hence we have proved the decomposition (31).

12

2) Exactly as in [D 85], decomposition (31) combined with assumption $(H6)$ concludes the proof of theorem 1.2. Thanks to $(H6)+(H4)$ and $\sqrt{n}$-consistency of the discretized estimator sequence $(\widetilde{\widetilde{\vartheta}}_n)_n$, we have an information with estimated parameter such that

$$I^{(\widetilde{\widetilde{\vartheta}}_n)} = I^{(\vartheta)} + o_{Q^\vartheta}(1) \quad \text{and} \quad (I^{(\widetilde{\widetilde{\vartheta}}_n)})^{-1} = (I^{(\vartheta)})^{-1} + o_{Q^\vartheta}(1) .$$

Exploiting (31) and $\sqrt{n}$-consistency of $(\widetilde{\widetilde{\vartheta}}_n)_n$, we have a score with estimated parameter satisfying

$$\Delta_n^{(\widetilde{\widetilde{\vartheta}}_n)} = \Delta_n^{(\vartheta)} - I^{(\vartheta)}[n^{1/2}(\widetilde{\widetilde{\vartheta}}_n - \vartheta)] + o_{Q^\vartheta}(1) .$$

Hence by definition of

$$\vartheta_n^* = \widetilde{\widetilde{\vartheta}}_n + n^{-1/2}(I^{(\widetilde{\widetilde{\vartheta}}_n)})^{-1} \Delta_n^{(\widetilde{\widetilde{\vartheta}}_n)}$$

we have immediately from (31)

$$\begin{aligned}
\sqrt{n}(\vartheta_n^* - \vartheta) &= \sqrt{n}(\widetilde{\widetilde{\vartheta}}_n - \vartheta) + (I^{(\widetilde{\widetilde{\vartheta}}_n)})^{-1} \Delta_n^{(\widetilde{\widetilde{\vartheta}}_n)} \\
&= \sqrt{n}(\widetilde{\widetilde{\vartheta}}_n - \vartheta) + (I^{(\widetilde{\widetilde{\vartheta}}_n)})^{-1} \left( \Delta_n^{(\vartheta)} - I^{(\vartheta)}[n^{1/2}(\widetilde{\widetilde{\vartheta}}_n - \vartheta)] + o_{Q^\vartheta}(1) \right) \\
&= (I^{(\vartheta)})^{-1} \Delta_n^{(\vartheta)} + o_{Q^\vartheta}(1) = Z_n^{(\vartheta)} + o_{Q^\vartheta}(1)
\end{aligned}$$

for the central sequence defined in (9). An application of (10) in corollary 1.1' finishes the proof. □

## 3. A minimum distance estimator sequence

We consider a sequence of minimum distance estimators (MDE) in the sense of Millar ([M 81], see also [Ku 04, Chapter 2.2], [H 08, Chapter 2.B+C]). We write $H$ for the Hilbert space $L^2([0,T], \mathcal{B}([0,T]), \lambda)$ with Lebesgue measure $\lambda$ and assume in this section

(33) $$\sigma^2(\cdot) \text{ is bounded and bounded away from } 0$$

together with $(H1)+(H3)$; slighly modified analogues to $(H2)$ and $(H4)$ will come in below.

To a path segment $\mathbb{X}_k = \left( \xi_{(k-1)T+s} \right)_{0 \le s \le T}$ we associate

$$\mathbb{Y}_k := \left( \xi_{(k-1)T+s} - \xi_{(k-1)T} - \int_{(k-1)T}^{(k-1)T+s} b(\xi_r)\, dr \right)_{0 \le s \le T}$$

and

$$\mathbb{V}_k := \left( \int_{(k-1)T}^{(k-1)T+s} \sigma(\xi_r)\, dW_r \right)_{0 \le s \le T}$$

for $k \geq 1$. Based on observation of a trajectory of the process (1) with unknown $\vartheta$ up to time $nT$, we define empirical quantities

$$\widehat{\Psi}_n \ := \ \frac{1}{n} \sum_{k=1}^{n} \mathbb{Y}_k \ , \quad n \geq 1$$

and compare these –thanks to $T$-periodicity of the signals– to

$$\Psi_\zeta \ := \ \left( \int_0^s dv \, S(\zeta, v) \right)_{0 \leq s \leq T}$$

for all possible values of the parameter $\zeta \in \Theta$. We have the representation

$$(34) \qquad \widehat{\Psi}_n \ = \ \Psi_\vartheta \ + \ \left( \frac{1}{n} \sum_{k=1}^{n} \int_{(k-1)T}^{(k-1)T+s} \sigma(\xi_r) \, dW_r \right)_{0 \leq s \leq T}$$

under $Q^\vartheta$, and prepare on some probability space a time-changed Brownian motion

$$(35) \qquad \mathbb{W} = (B \circ \Phi(s))_{0 \leq s \leq T} \quad , \quad \Phi(s) \ := \ \int_0^s dv \left[ \mu^{(\vartheta)} P_{0,v}^{(\vartheta)} \, \sigma^2 \right] \ , \quad 0 \leq s \leq T \ .$$

With the usual conventions, our MDE sequence for the unknown parameter $\vartheta \in \Theta$ will be

$$(36) \qquad \widetilde{\vartheta}_n \ := \ \operatorname*{arginf}_{\zeta \in \Theta} \| \widehat{\Psi}_n - \Psi_\zeta \|_H \ , \quad n \geq 1 \ .$$

We strengthen the identifiability condition $(H2)$ to

$$(37) \qquad \inf_{\zeta \in \Theta, |\zeta - \vartheta| > \varepsilon} \| \Psi_\zeta - \Psi_\vartheta \|_H \ > \ 0$$

for all $\vartheta \in \Theta$, and have by means of theorem A in section 2 a weak law of large numbers

$$(38) \qquad \| \widehat{\Psi}_n - \Psi_\vartheta \|_H \ \longrightarrow \ 0 \quad \text{in probability under } Q^\vartheta \text{ as } n \to \infty$$

together with an integrability property

$$(39) \qquad \sup_{n \geq 1} E_\vartheta \left( n \left( \widehat{\Psi}_n(s) - \Psi_\vartheta(s) \right)^2 \right) \ \leq \ C \, s \quad , \quad \lim_{n \to \infty} E_\vartheta \left( n \left( \widehat{\Psi}_n(s) - \Psi_\vartheta(s) \right)^2 \right) \ = \ \Phi(s)$$

for every $0 \leq s \leq T$. By finite-dimensional convergence combined with (39), the processes (34) under $Q^\vartheta$ converge weakly in the path space $H$ to the Gaussian process $\mathbb{W}$ defined in (35), see e.g. Cremers and Kadelka [CK 86]. This implies ([M 81], [Ku 04], or [H 08, 2.13]) that the MDE sequence (36) converges in $Q^\vartheta$-probability to the true value $\vartheta$ and is tight at rate $\sqrt{n}$, for all $\vartheta \in \Theta$. Thus we dispose of a *preliminary estimator sequence for $\vartheta$ which is tight at the required rate*. Already at this stage, we can apply theorem 1.2 and obtain an easy to calculate modified estimator sequence which is efficient in the sense of the local asymptotic minimax bound given in corollary 1.1'. If in addition, for all $\vartheta \in \Theta$, the components $\dot{S}_j(\vartheta, \cdot)$ of the derivative in $(H3)$ satisfy

$$(40) \qquad D_j \Psi_\vartheta = \left( \int_0^s \dot{S}_j(\vartheta, v) \, dv \right)_{0 \leq s \leq T} , \ 1 \leq j \leq d \ , \quad \text{are linearly independent in } H$$

14

there is an explicit limit law for rescaled estimation errors ([M 81], [Ku 04], or [H 08, 2.23]) of the MDE sequence. Summarizing we have

**3.1 Theorem:** Under $(H1)$–$(H6)$ and $(33)+(37)+(40)$, we have for all $\vartheta \in \Theta$ weak convergence

$$\mathcal{L}\left(\sqrt{n}(\widetilde{\vartheta}_n - \vartheta) \mid Q^\vartheta\right) \quad \longrightarrow \quad \mathcal{N}\left(0, \Lambda_\vartheta^{-1} \Xi_\vartheta \Lambda_\vartheta^{-1}\right)$$

in $\mathbb{R}^d$ as $n \to \infty$, where $\Lambda_\vartheta$ denotes the $d \times d$-matrix with entries $\int_0^T D_i \Psi_\vartheta(s) D_j \Psi_\vartheta(s) ds$, and $\Xi_\vartheta$ the $d \times d$-matrix with entries $\int_0^T \int_0^T D_i \Psi_\vartheta(s_1) \Phi_\vartheta(s_1 \wedge s_2) D_j \Psi_\vartheta(s_2) ds_1 ds_2$ for $\Phi_\vartheta$ defined in (35). One-step modification according to theorem 1.2 transforms the MDE sequence $(\widetilde{\vartheta}_n)_n$ into an estimator sequence $(\vartheta_n^*)_n$ which is asymptotically efficient at all points $\vartheta \in \Theta$ in the sense of corollary 1.1'.

Other estimator sequences do exist which are efficient in the sense of the local asymptotic minimax bound of corollary 1.1'.

**3.2 Remark :** In addition to (33), let $\Theta$ be bounded and assume that

(41) $\qquad \underline{C}|\zeta' - \zeta|^2 \quad \leq \quad \int_0^T (S(\zeta', t) - S(\zeta, t))^2 \, dt \quad \leq \quad \overline{C}|\zeta' - \zeta|^2 \quad \text{for all } \zeta, \zeta' \in \Theta$

for suitable constants $0 < \underline{C}, \overline{C} < \infty$ (compare (41) to (11)+(12)). Then maximum likelihood (MLE)

$$\vartheta_n^{(*,1)} = \operatorname*{argmax}_{\zeta \in \overline{\Theta}} L_{nT}^{\zeta/\zeta_0} \quad, \quad n \in \mathbb{N}$$

(with $\zeta_0 \in \Theta$ some fixed point, and $\overline{\Theta}$ the closure of $\Theta$) and Bayes estimator (BE) sequences

$$\vartheta_n^{(*,2)} := \frac{\int_\Theta \zeta \, L_{nT}^{\zeta/\zeta_0} \, d\zeta}{\int_\Theta L_{nT}^{\zeta/\zeta_0} \, d\zeta} \quad, \quad n \in \mathbb{N}$$

will be efficient in the sense of corollary 1.1'.

The main ideas can be found in Ibragimov and Khasminskii [IH 81], LeCam and Yang [LY 90], Kutoyants [K 04]. They have to be adapted to the time inhomogeneous setting with $T$-periodic semigroup which is the case here. Without going into details, we mention that under $(33)+(41)$ there are bounds for Hellinger type distances and affinities of form

(42) $\qquad E_\zeta\left(\sup_{0 \leq s \leq t} L_t^{\zeta'/\zeta}\right) \quad \leq \quad c_0 + c_1\left(\left\lfloor \frac{t}{T} \right\rfloor + 1\right)|\zeta' - \zeta|^2$

(43) $\qquad E_\zeta\left(\left[1 - \left(L_t^{\zeta'/\zeta}\right)^{1/2}\right]^2\right) \quad \leq \quad \sum_{j=1}^3 c_j\left(\left\lfloor \frac{t}{T} \right\rfloor + 1\right)^j |\zeta' - \zeta|^{2j}$

$$(44) \qquad E_\zeta \left( \left[ 1 - \left( L_t^{\zeta'/\zeta} \right)^{1/4} \right]^4 \right) \;\; \leq \;\; \sum_{j=2}^{5} c_j \left( \left\lfloor \frac{t}{T} \right\rfloor + 1 \right)^j |\zeta' - \zeta|^{2j}$$

$$(45) \qquad E_\zeta \left( \left[ L_t^{\zeta'/\zeta} \right]^{1/2} \right) \;\; \leq \;\; \exp \left\{ - k \left\lfloor \frac{t}{T} \right\rfloor |\zeta' - \zeta|^2 \right\}$$

with positive constants $c_j, k$ which do not depend on $\zeta, \zeta' \in \Theta$ or on $t \geq 0$: based on ideas from the references above, a proof similiar to [HK 10, proof of lemma 5.1] goes through with

$$\delta_s := \frac{S(\zeta', s) - S(\zeta, s)}{\sigma(\eta_s)} \quad , \quad d(s) := |S(\zeta', s) - S(\zeta, s)|^2$$

$$\underline{c}\, d(s) \;\leq\; \delta_s^2 \;\leq\; \overline{c}\, d(s) \quad , \quad \underline{C} \left\lfloor \frac{t}{T} \right\rfloor |\zeta' - \zeta|^2 \;\leq\; \int_0^t d(s) ds \;\leq\; \overline{C} \left( \left\lfloor \frac{t}{T} \right\rfloor + 1 \right) |\zeta' - \zeta|^2$$

where we exploit (33)+(41), with suitable constants. Combining (43)–(45) with convergence of experiments according to theorem 1.1, one follows the approach of Ibragimov and Khasminskii [IH 81] towards convergence of MLE and BE together with their moments of arbitary order. One can also use LeCam's 'Third lemma' to consider these estimator sequences under contiguous alternatives. For the sequence of local models at a reference point $\vartheta$ defined according to (7) in theorem 1.1, let $\widetilde{\mathcal{E}} = \{ \widetilde{P}_u : u \in I\!\!R^d \}$ denote the Gaussian limit experiment with central statistics $u^*$. Then $u^*$ is the MLE in the limit experiment, and also the BE ([IH 81, p. 180], [K 04, p. 134]). Thus MLE and BE in the Gaussian limit experiment $\widetilde{\mathcal{E}}$ are equivariant estimators, the law of their error under arbitrary $u \in I\!\!R$ being equal to $\mathcal{L}(u^* | \widetilde{P}_0)$. For both sequences, we thus obtain

$$(46) \qquad \lim_{n \to \infty} \sup_{|u| \leq C} \left| E_{\vartheta + n^{-1/2} u} \left( \ell \left( n^{1/2} (\vartheta_{nT}^{(*,i)} - (\vartheta + n^{-1/2} u)) \right) \right) - E_{\widetilde{P}_0} \left( \ell \left( u^* \right) \right) \right| \;\; = \;\; 0$$

for arbitrary loss functions which are continuous subconvex with polynomial majorant, and for arbitrarily large constants $C$. Combined with corollary 1.1', (46) is efficiency of both the MLE and the BE sequence, simultaneously for a large class of loss functions, in the sense of the local asymptotic minimax bound of corollary 1.1'.

# References

[ADR 69]   Azéma, J., Duflo, M., Revuz, D.: Mesures invariantes des processus de Markov récurrents. Séminaire de Probabilités III, Lecture Notes in Mathematics **88**, 24–33. Springer 1969.

[CLM 06]   Castillo, I., Lévy-Leduc, C., Matias, C.: Exact adaptive estimation of the shape of a periodic function with unknown period corrupted by white noise. Mathe. Meth. Statist. **15**, 1–30 (2006).

[CK 86]   Cremers, H., Kadelka, D.: On weak convergence of integral functions of stochastic processes with application to processes taking paths in $L_p^E$. Stoch. Proc. Applic. **21**, 305–317 (1986).

[D 85]   Davies, R.: Asymptotic inference when the amount of information is random. In: Le Cam, L., Olshen, R. (Eds): Proc. of the Berkeley Symposium in honour of J. Neyman and J. Kiefer. Vol. II. Wadsworth 1985.

[H 70]   Hájek, J.: A characterization theorem of limiting distributions for regular estimators. Zeitschr. Wahrscheinlichkeitstheor. Verw. Geb. **14**, 323–330, 1970.

[H 08]   Höpfner, R.: Asymptotische Statistik. Manuscript in progress. Mainz 2008. Under `http://www.mathematik.uni-mainz.de/∼hoepfner`

[HK 10]   Höpfner, R., Kutoyants, Y.: Estimating discontinuous periodic signals in a time inhomogeneous diffusion. Submitted to Statist. Inference Stoch. Proc., `arXiv:0903.5061`.

[IH 81]   Ibragimov, I., Has'minskii, R.: Statistical estimation. Springer 1981.

[JS 87]   Jacod, J., Shiryaev, A.: Limit theorems for stochastic processes. Springer 1987.

[KS 91]   Karatzas, J., Shreve, S.: Brownian motion and stochastic calculus. 2nd ed. Springer 1991.

[K 04]   Kutoyants, Y.: Statistical inference for ergodic diffusion processes. Springer 2004.

[L 68]   Le Cam, L.: Théorie asymptotique de la décision statistique. Montréal 1969.

[LY 90]   Le Cam, L., Yang, G.: Asymptotics in statistics. Some basic concepts. Springer 1990. (2nd Ed. Springer 2002).

[LS 81]   Liptser, R., Shiryaev, A.: Statistics of random processes. Vols. I+II, Springer 1981, 2nd Ed. 2001.

[MT 93]   Meyn, S., Tweedie, R.: Markov chains and stochastic stability. Springer 1993.

[M 83]   Millar, P.: The minimax principle in in asymptotic statistical theory. In: P. Hennequin (Ed.), Ecole d'Eté de Probabilités de St. Flour XI 1981. Lect. Notes in Math. 976, Springer 1983.

[R 75]   Revuz, D.: Markov chains. North Holland 1975.

[RY 91]   Revuz, D., Yor, M.: Continuous martingales and Brownian motion. Springer 1991.

[S 65]   Skorokhod, A.: Studies in the theory of random processes. Addison-Wesley 1965.

Yury A. Kutoyants

Laboratoire de Statistique et Processus, Université du Maine, F–72085 Le Mans Cedex 9

`kutoyants@univ-lemans.fr`

`http://www.univ-lemans.fr/sciences/statist/pages_persos/kuto.html`

Reinhard Höpfner

Institut für Mathematik, Universität Mainz, D–55099 Mainz

`hoepfner@mathematik.uni-mainz.de`

`http://www.mathematik.uni-mainz.de/∼hoepfner`